Developing an Open-Source Database and Pythonic Toolset for Quantifying Lacustrine Sedimentation using Publicly Available Data

Jake Gearon[1] and John Franey[2]
University of Texas, Jackson School of Geosciences; *Dept. Of Geological Sciences[1] and Bureau of Economic Geology[2]*

TEXAS Geosciences
The University of Texas at Austin
Jackson School of Geosciences

Introduction

There are many publicly available lake-level data sets in varying states of accessibility, upkeep, and accuracy (USGS, HydroWeb, Copernicus Global Land Service, NASA, NOAA, and various publications)

The collating of data from disparate sources, or data-wrangling, is often the barrier to entry for many scientific projects. Lake water level data particularly suffers from lack of ease of access as it is distributed haphazardly across federal, state, and academic databases.

We aim to provide an open-source, well documented, scalable, modular, and regularly updated database of lake water levels and a python package of lacustrine analysis tools for use by researchers across disciplines and institutions.

USDA United States Department of Agriculture
Foreign Agricultural Service
Global Reservoirs and Lakes Monitor (G-REALM)

Theia HydroWeb

USGS science for a changing world NWIS

+ others

amazon RDS

Amazon API Gateway → Researcher

python GitHub

Open-Source Lake Level Database

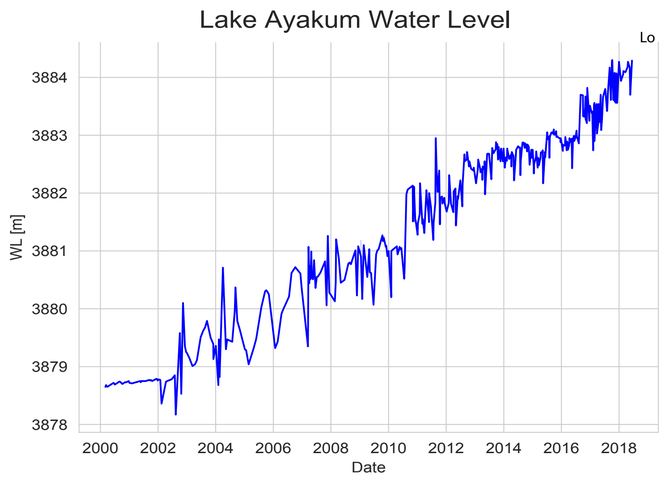
There are far more databases administered at the municipal and state level than is possible to include in this project, therefore a user-friendly, documented workflow for easily adding additional datasets via forking on GitHub will be a key component of this project.

Additionally, users will be able to modify or add functionalities to the "LakePy" python package for further flexibility and power.

open source initiative

LakePy Toolbox

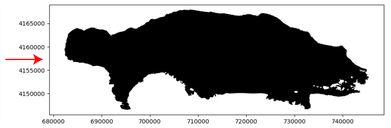
We aim to build a Python package that natively interfaces with the lake-level database—"LakePy"— which is the preferred method of access. "LakePy" will host a basic set of lacustrine & remote sensing analysis methods including time-series generation, shoreline identification and extraction. Any user will be able to submit additional tools/analysis options via forking the project.



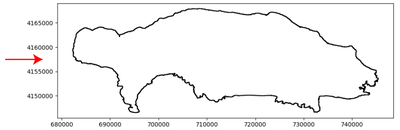
Net Sedimentation Analysis

"LakePy" along with a set of standard remote sensing analytical capabilities will include an original sedimentation analysis toolkit to assess the "net sedimentation state" of a lake over a user-defined time period. Below is a schematic example of the analysis.

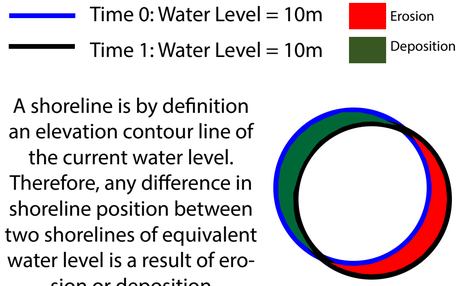
Satellite Imagery (LANDSAT, Sentinel-2)



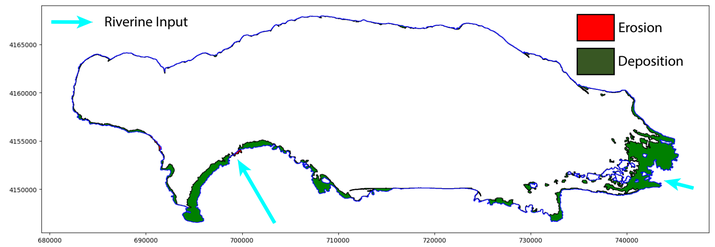
Modified Normalized Difference Water Index (MNDWI)



Morphological Mathematics (Geleynse et al., 2012)



A shoreline is by definition an elevation contour line of the current water level. Therefore, any difference in shoreline position between two shorelines of equivalent water level is a result of erosion or deposition



Significance & Future Work

Removing hurdles between researchers and available data is imperative in an increasingly data-rich world. Many researchers have immense subject-matter expertise regarding the biological, hydrological, and sedimentological nuances of lacustrine environments but little or no experience in automated querying, data cleaning/wrangling, or programming.

Acknowledgments

This work is based on funding provided by the ESIP Lab with support from the National Aeronautics and Space Administration (NASA), National Oceanic and Atmospheric Administration (NOAA) and the United States Geologic Survey (USGS). The authors would also like to acknowledge the Dynamic Stratigraphy Group at UT Austin